



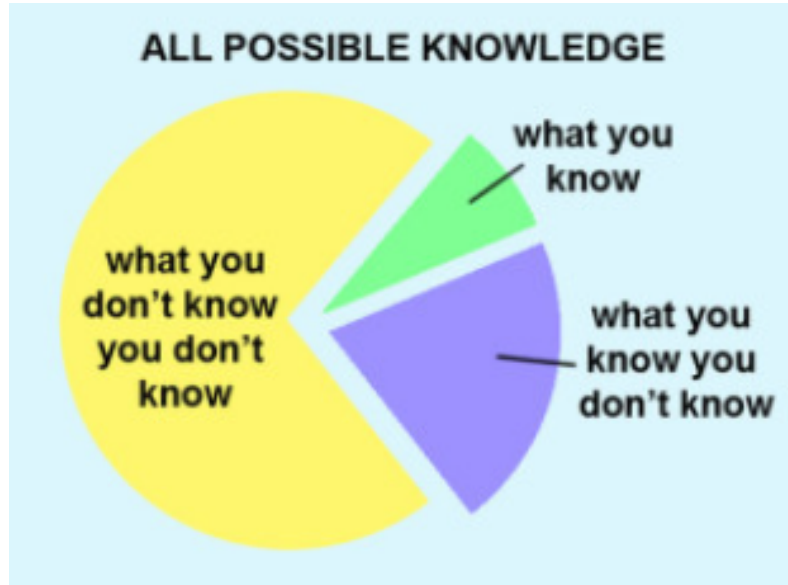
Value-Added Validation:

Regulatory Review of Data Submission Packages

Mary Anne Potok
26 September 2019 – PhUSE SDE



Submission Data Packages



- Largest and most technically challenging files
- Root of all importance for submission review
- Least able to be understood by Regulatory

Typical eSubmission Data Package

- Tabulation datasets
- Tabulation data definition
- Study Data Reviewer's Guide (CSDRG or NSDRG)
- Annotated CRF
- Analysis datasets
- Analysis data definition
- Analysis Data Reviewer's Guide (ADRG)
- Analysis program files



Permitted File Types

- Only certain file types are allowed in FDA submissions, and only in certain eCTD sections

File Type	Format Name	Location Allowed
.pdf	Portable Document Format	M1-M5
.doc, .docx	Microsoft Word document	M1.14, 1.16
.txt	Text file	M3-M5
.xls, .xlsx	Microsoft Excel document	M3-M5 (also include PDF)
.css	Cascading Style Sheets	M1.15, M4-M5
.xml	Extensible Markup Language	M1.14, M4-M5
.mpeg, .mp3, mp4, .wav	Audio and Video files	M1.15
.xpt	SAS Transport file	M3-M5

Data Package Conversion

- Programming and Stats deliverables cannot be submitted without conversion to acceptable file formats for the eCTD

Deliverable Type	File Format Required	File Format Produced
Tables, Listings, Figures	.pdf	.rtf
Datasets	.xpt	.sas7bdat
SAS Codes	.txt	.sas
Annotated CRF	No annotations allowed	PDF with annotations

'Inside' vs 'Outside' Review of Data Packages

Programming Review

- Review of contents
- Meet all Agency requirements
- Complies with data standards



Reg Ops Review

- No files missing and all can be opened
- File type (extension) is correct
- Review of file tags



Regulatory Operations QC of Data Packages

Data Packages (at delivery)

Datasets converted to required transport format (.xpt)

Data definitions provided for each study

TS Domain provided for each study

Program files, aCRF meet FDA format requirements

Submission Output (after publishing)

All files are present, have correct study number, and in correct eCTD section

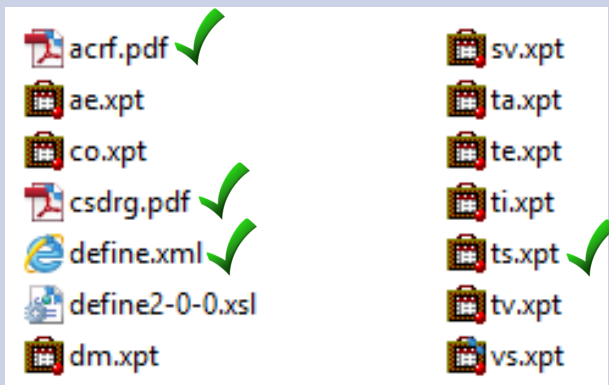
Data file size in published output matches source file

Datasets organized by appropriate ICH file tag

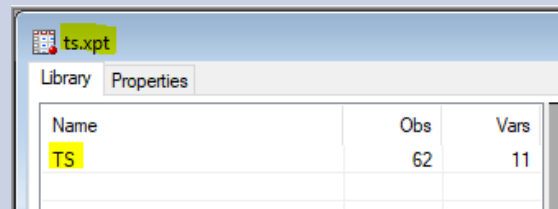
eCTD Validator Report free of errors

Regulatory Operations QC of Data Packages

Data Packages (at delivery)



Submission Output (after publishing)



A screenshot of the 'ts.xpt' file properties window. The 'Properties' tab is selected, showing the following data:

Name	Obs	Vars
TS	62	11

	STUDYID	DOMAIN	TSSEQ	TSPARMCD
1	Study ID	TS		1 ACTSUB
2		TS		1 ADAPT
3		TS		1 ADDON
4		TS		1 AGEMAX

Study Tagging File (stf.xml)

- Lists every file related to that single study
- Similar to a TOC for each individual study

Study Tagging File (STF)

Study Title:	STF - A Double-Blind, Double-Dummy, Randomized, Positive and Placebo Controlled, Cross-Over Study of the Effects of Intervals in Healthy Subjects		
Study ID:	Study ID	TQTC	Study ID
Category:			
Study Content:			
Document Title:	Study ID	Protocol	Relative Filename: m5/53-clin-stud-rep/534-rep-human-pd-stud/5341-healthy-subj-pd-stud-rep-protocol.pdf
			<u>[ich] protocol-or-amendment</u>
Document Title:	Study ID	CSR	Relative Filename: m5/53-clin-stud-rep/534-rep-human-pd-stud/5341-healthy-subj-pd-stud-rep-report-body-1.pdf
			<u>[ich] study-report-body</u>
Document Title:	Study ID	Narratives	Relative Filename: m5/53-clin-stud-rep/534-rep-human-pd-stud/5341-healthy-subj-pd-stud-rep-report-body-2.pdf
			<u>[ich] study-report-body</u>
Document Title:	Highlights of Clinical Pharmacology		Relative Filename: m5/53-clin-stud-rep/534-rep-human-pd-stud/5341-healthy-subj-pd-stud-rep-report-body-3.pdf
			<u>[ich] study-report-body</u>
Document Title:	define		Relative Filename: m5/datasets/Study ID/tabulations/legacy/define.pdf
			<u>[us] data-tabulation-data-definition</u>
Document Title:	define-ecg		Relative Filename: m5/datasets/Study ID/tabulations/legacy/define-ecg.pdf
			<u>[us] data-tabulation-data-definition</u>
Document Title:	AE		Relative Filename: m5/datasets/Study ID/tabulations/legacy/ae.xpt
			<u>[us] data-tabulation-data set-legacy</u>
Document Title:	CM		Relative Filename: m5/datasets/Study ID/tabulations/legacy/cm.xpt

Submission Building with Study Tags

Outline

Filter <none>

0000 6 - 5 0 1

- 0000
 - 1 Administrative Information and Prescribing Information
 - 2 Common Technical Document Summaries
 - 3 Quality
 - 4 Nonclinical Study Reports
 - 5 Clinical Study Reports
 - 5.2 Tabular Listing of all Clinical Studies
 - 5.3 Clinical Study Reports
 - 5.3.1 Reports of Biopharmaceutic Studies
 - 5.3.2 Reports of Studies Pertinent to Pharmacokinetics using Human Bi
 - 5.3.3 Reports of Human Pharmacokinetic (PK) Studies
 - 5.3.4 Reports of Human Pharmacodynamic (PD) Studies
 - 5.3.5 Reports of Efficacy and Safety Studies
 - 5.3.5.1 Study Reports of Controlled Clinical Studies Pertinent to the C
 - <Insert Study Title>
 - Synopsis
 - Report Body
 - 16.1 Appendices
 - 16.2 Appendices
 - CRFs
 - Datasets
 - Analysis Data Sets (Legacy)
 - Analysis Data Sets (ADaM)
 - Data Definitions
 - Program File
 - Datasets
 - Analysis datasets (ADaM)
 - Data Tabulation Datasets (Legacy)
 - Data Tabulation Datasets (SDTM)

Properties

General

Structure Number

Title

Program File

eSubmission Code

Analysis Datasets (ADaM) - Program File

Node Title Option

Configuration

Output Media

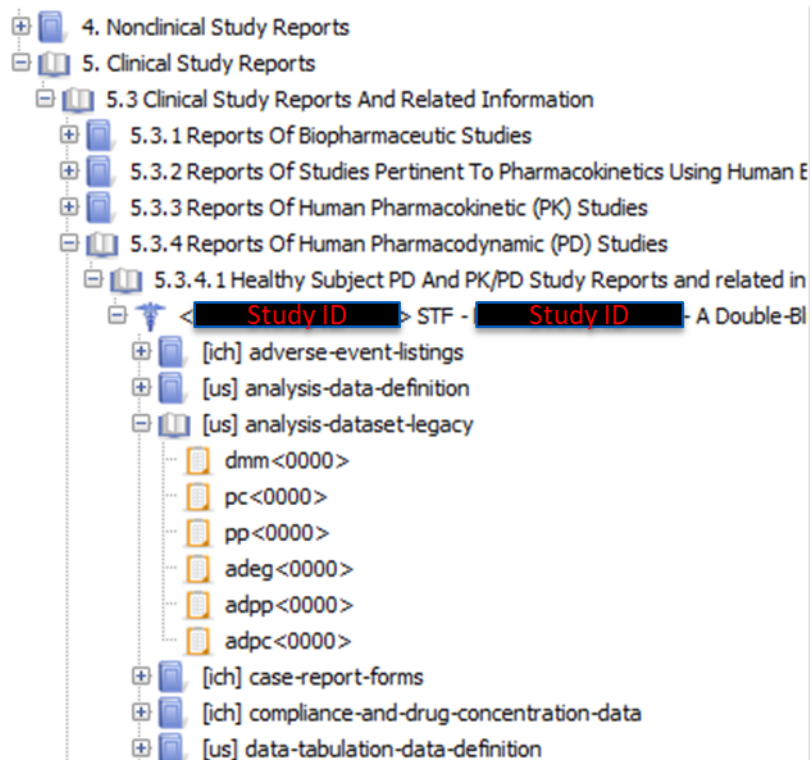
Properties Find Hyperlinks

Node Content

- Analysis Datasets (Legacy)
- Analysis Datasets (Legacy) - Split
- Analysis Datasets (Legacy) - Data Definition
- Analysis Datasets (Legacy) - Annotated CRF
- Analysis Datasets (Legacy) - Program File
- Analysis Datasets (ADaM)
- Analysis Datasets (ADaM) - Split
- Analysis Datasets (ADaM) - Data Definition
- Analysis Datasets (ADaM) - Annotated CRF
- Analysis Datasets (ADaM) - Program File
- Tabulation Datasets
- Tabulation Datasets (Legacy)
- Tabulation Datasets (Legacy) - Split
- Tabulation Datasets (Legacy) - Data Definition
- Tabulation Datasets (Legacy) - Annotated CRF
- Tabulation Datasets (SDTM)
- Tabulation Datasets (SDTM) - Split
- Tabulation Datasets (SDTM) - Data Definition
- Tabulation Datasets (SDTM) - Annotated CRF
- Tabulation Datasets (SEND)
- Tabulation Datasets (SEND) - Data Definition
- Miscellaneous Datasets
- Miscellaneous Datasets - Data Definition
- Miscellaneous Datasets - Annotated CRF
- Data Listing Datasets (BIMO)
- Data Listing Datasets (BIMO) - Data Definition

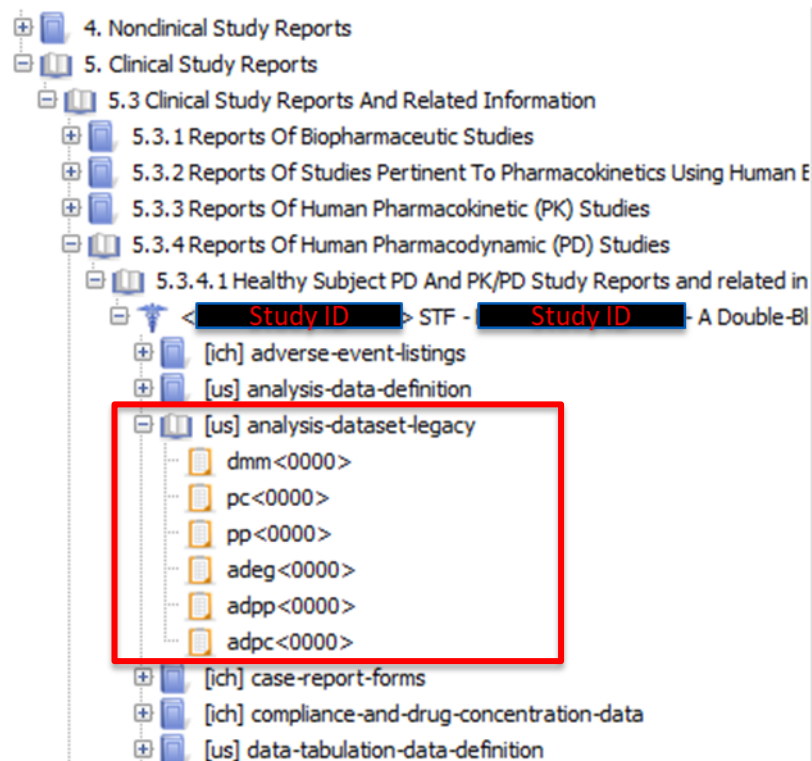
Study Tagging File (stf.xml)

- Choosing the correct file tag is a manual process
- Validation will not identify tagging errors



Study Tagging File (stf.xml)

- Choosing the correct file tag is a manual process
- Validation will not identify tagging errors



'Inside' vs 'Outside' Validation

Pinnacle 21 Looks *Inside*

Issue Summary			
Publisher ID	Message	Severity	Found
FDAC036	Variable length is too long for actual data	Error	17
FDAC128	Missing AESTDY variable, when AESTDTC variable is present	Error	1
FDAC132	Missing AEENDY variable, when AEENDTC variable is present	Error	1
FDAC021	FDA Expected variable EPOCH not found	Warning	1
FDAC055	Permissible variable with missing value for all records	Warning	6
FDAC036	Variable length is too long for actual data	Error	16
FDAC128	Missing CMSTDY variable, when CMSTDTC variable is present	Error	1
FDAC132	Missing CMENDY variable, when CMENDTC variable is present	Error	1
FDAC341	CMDOSU value not found in 'Unit' extensible codelist	Warning	2
FDAC021	FDA Expected variable EPOCH not found	Warning	1
FDAC055	Permissible variable with missing value for all records	Warning	1

eCTD Validator Looks *Outside*

FDA Categories / Study Data / Study Data Standards

1734

High(2)

A Trial Summary (TS) dataset must be present for each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2

Missing files

- Missing ts.xpt for study id
- Missing ts.xpt for study id

1736

High(3)

DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4. DM dataset, ADSL dataset, define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2

Missing file(s)

- Missing define.xml for study id
- Missing define.xml for study id
- Missing define.xml for study id

Critical Steps in Data Package Review

	CDISC Validator	eCTD Validator	Manual Submission Review
Domains and domain names	Yes	No	No
Variables	Yes	No	No
Controlled terminology	Yes	No	No
Fitness of data for review	Yes	No	No
File types	Yes	Yes	Yes
Active hyperlink	Yes	Yes	Yes
Data requirements per eCTD section	No	Yes	Yes
Reviewer's Guide available	Yes	Yes	Yes
Reviewer's Guide meets PDF specs	No	Yes	Yes

eCTD Validation Criteria

- Technical Rejection Criteria for Study Data
- Specifications for eCTD Validation Criteria

FDA Categories / Study Data / Study Data Standards		
1734	OK	A Trial Summary (TS) dataset must be present for each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1735	OK	The correct STF file-tags must be used for all standardized datasets in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1736	OK	DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4. DM dataset, ADSL dataset, define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1737	OK	For each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2, no more than one dataset of the same name should be submitted as new

eCTD Validation Criteria

- Technical Rejection Criteria for Study Data
- Specifications for eCTD Validation Criteria

FDA Categories / Study Data / Study Data Standards		
1734	OK	A Trial Summary (TS) dataset must be present for each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1735	OK	The correct STF file-tags must be used for all standardized datasets in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1736	OK	DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4. DM dataset, ADSL dataset, define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
1737	OK	For each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2, no more than one dataset of the same name should be submitted as new

Resources for Regulatory Operations

Study Data Technical Conformance Guide

- SDTM, ADaM, SEND standards
- Sizing, naming, organization

eCTD Technical Conformance Guide

- Metadata and study tagging

Technical Rejection Criteria for Study Data

- 4 errors identified using eCTD Validator
- CTD sections that require data

Specifications for eCTD Validation Criteria

- Explanation and remediation for eCTD validation errors

Resources for Regulatory Operations

Study Data Technical Conformance Guide

- SDTM, ADaM, SEND standards
- Sizing, naming, organization

eCTD Technical Conformance Guide

- Metadata and study tagging

Technical Rejection Criteria for Study Data

- 4 errors identified using eCTD Validator
- CTD sections that require data

Specifications for eCTD Validation Criteria

- Explanation and remediation for eCTD validation errors

Resources for Regulatory Operations

Data Teams!

Resources for Regulatory Operations

Data Teams!

- Equally responsible to make sure the largest, most technical, and most critical submission components are accessible to reviewers
- Thorough Quality Control checks of content ***and*** assembly
- Work together to mitigate risks of Refusal to File and Technical Rejection

Questions?